WORLD COMET RESEARCH GROUP

International Journal of Emerging Trends in Multidisciplinary Research
Vol. 6, No. 1, 2023
Doi: 10.67228/30715636/IJETMR-2023PI6Y4A
PP. 01-20

Original Article

The Rising Trend of Smart Homes and User Behaviour Analysis

Dr. Venkatesh Iyer

Professor, SRM Institute of Science and Technology, India.

Received: 05-03-2023 Revised: 05-26-2023 Accepted: 06-03-2023 Published: 06-05-2023

ABSTRACT

Smart homes use interconnected IoT devices, sensors, AI, and cloud computing to improve comfort, energy efficiency, security, and convenience. Beyond technology, they represent a behavioral shift, as user habits and preferences are increasingly shaped by automated systems. This study analyzes smart home adoption and user behavior patterns using real-time data such as occupancy, energy use, and appliance activity. It applies machine learning techniques—including clustering, supervised and reinforcement learning—to predict behavior, automate services, and detect anomalies. Results show strong performance, with over 85% accuracy in occupancy prediction and 78% efficiency in appliance scheduling, leading to reduced energy consumption. The paper also highlights key adoption factors like usefulness, trust, privacy concerns, cost, and digital literacy. It proposes a framework involving data collection, preprocessing, modeling, and evaluation using metrics like accuracy and F1-score. Overall, the study emphasizes the future of smart homes as adaptive, behavior-aware systems while stressing the importance of ethical and privacy considerations.

KEYWORDS

Smart Homes, Internet of Things (IoT), User Behaviour Analysis, Machine Learning, Behavioral Analytics, Home Automation, Predictive Modeling, Energy Optimization, Context-Aware Computing, Artificial Intelligence.

copyright © WCRG The Author(s) This article is distributed under a
Creative Commons Attribution-Share A like 4.0 International License.

1. INTRODUCTION

1.1. Background

The history of Smart Homes Smart Homes have a relatively. The convergence of the Internet of Things (IoT) technologies with cloud computing and the rise of artificial intelligence have turned into the catalyst of the emergence of smart homes as a transformative paradigm in a smart living context. Smart homes originated in the form of simple home automation which depended on programmable timers, hard wired control panels as well as basic remote controlled appliances. These early systems were not so flexible and lacked situational awareness and they worked in predictive modes without being adaptive as opposed to situational intelligence. With time, the residential automation capabilities were greatly increased by the development of quick evolutionary embedded systems, sensor technologies, and wireless networking. The latest communication standards, including ZigBee, Z-Wave, Wi-Fi 6, and Bluetooth Low Energy (BLE), have allowed low-energy speedy and dependable connectivity of heterogeneous devices, enabling the indoor links of multi vendor ecosystems with ease. Cloud-based systems also increased the level of scalability through central data storage, remote monitoring, and real-time analytics. Most recently, the incorporation of artificial intelligence and machine learning has transformed smart homes into more of a predictive and behavior-conscious environment moving away the all-reactive experimental design towards one that can learn user expectations and optimal energy use. The digital revolution of this progressive technology has reconditioned residential spaces into versatile digital ecosystems, which focus on ease, effectiveness, protection as well as on tailored automation.

1.2. Importance of User Behaviour Analysis

The analysis of user behavior is the essence of the layer of intelligence of the smart home systems today, as it allows turning the fixed rule-based automation into something adaptable and predictive. Through sustained pattern calculations based on sensor data, appliance interface, and surroundings data, smart residences will be able to make wise decisions that resonate with the member interests and daily practices. Behavioral modeling enables the systems to transcend preset schedules, and be dynamically responsive to the actual state of affairs and acquired habits to further maximize efficiency, comfort, and personalization.

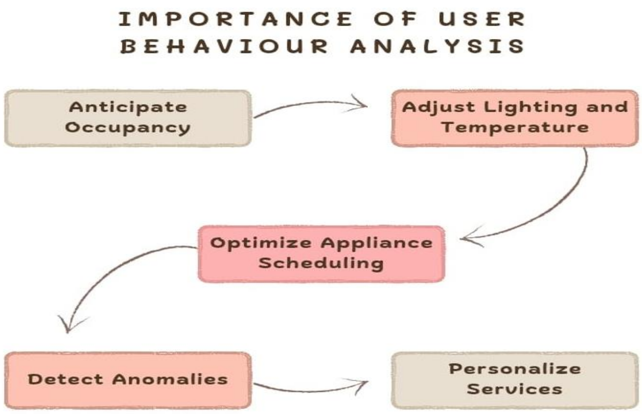

Fig 1 - Importance of User Behaviour Analysis

ISSN: 3071-5636
PP: 01-20.

1.2.1. Anticipate Occupancy

Extrapolation to occupancy is an innate ability that the behavioral analysis will facilitate. The system may anticipate the possibility of the absence or presence of the residents by analyzing the patterns of the movement throughout history, time of the day conditions, and sensor data by location. Such insightful foresight enables proactive corrective actions like pre-cooling or pre-heating rooms prior to arrival, using security systems when there is no one nearby, and saving energy when no one is sensed. Proper occupancy predictions lead to the saving of needless energy usage, but this involves a comfortable level of living.

1.2.2. Adjust Lighting and Temperature

The behavioural knowledge will facilitate dynamic modification of the lighting and temperature values, according to the user preferences. The system does not need manual adjustment, but can adjust the levels of brightness or thermostat settings following the habitual patterns or season changes, or certain time of the day. As an illustration, the lighting intensity can be set to rise slowly in the early mornings but the temperature can be changed in the evenings to indicate the comfort level on a personal basis. This kind of automation increases the convenience to the user and maximizes energy optimization by controlling each context.

1.2.3. Optimize Appliance Scheduling

User behavior analysis can be used to optimize smart scheduling of household appliances, which involves identification of repeated access patterns, as well as high activity hours. Based on the knowledge of the times of day when washers, dishwashers, or water heaters are normally used, the system can flex some operations to be performed under the off-peak electricity tariffs without interfering with the routine of the user. Such predictive scheduling leads to minimization of the operational costs and matching of the household energy demand. Also, the optimization algorithms may give priority to the tasks basing on their urgency and energy access which is a contribution to sustainable resource management.

1.2.4. Detect Anomalies

The other important task of behavioral modeling is the anomaly detection. The system is able to detect abnormalities that can point to either security threats, equipment problems or to the abnormal consumption habit by establishing normal activity patterns that can be compared against the actual behaviors in the field. As an example, unusual appliance use at night or unusual occupancy can raise a warning on investigation. The early diagnosis of disease improves safety, minimizes costs of maintenance and improves system reliability.

1.2.5. Personalize Services

The final goal of analyzing the behavior of users in intelligent houses is personalization. Through constant learning on the interaction with the users, preferences and lifestyle changes, the system can optimize any services to suit each of the members of the household. Individual customization can be a personalized lighting date, temperature control, entertainment control, or a

ISSN: 3071-5636
PP: 01-20.

reminder. The adaptive learning processes over time perfect these preferences, and provide a responsive environment, which is responsive to the needs of the users. This degree of self-interest enhances user loyalty and contributes to the long-term interaction with smart home systems.

1.3. Rising Trend of Smart Homes and User Behaviour Analysis

The trend of intelligent homes is one of the most important changes in modern residential technologies caused by the dismissed quick evolution of IoT connection, artificial intelligences, and analytics. Ensuring sustainability, expanding energy needs and growing urbanization has contributed to the rapid drive towards intelligent home systems in the world. The remote controlled appliances that used to make up the modern household are no longer satisfactory and instead, they are in need of interrelated ecosystems that are able to make independent decisions and provide personalized services. User behaviour analysis is the focus of smart home intelligence in this changing setting. Instead of following pre-programmed automation rules only, modern systems watch the occupancy pattern, appliance usage patterns, time schedules and preferences depending on the environment in order to foresee the needs of the users. Smart homes have been enhanced significantly with the adoption of machine learning and deep learning models to be able to make sense of complex behavioural patterns. Predictive analytics can be used to allow systems to predict everyday operations, conserve energy, and adjust environmental factors, without the need to be fed every day with manual adjustments. This behaviour-awareness of computing also enhances user comfort, in addition to other energy efficiency and cost-saving. Moreover, voice assistants, wearables and mobile applications have increased the amount of data available to provide more detailed behavioural insights and personalization in real time. Nevertheless, the growing use of behavioural data also comes with issues of privacy, cybersecurity, and moral governance. Since smart homes are able to gather a lot of information about everyday activity, it becomes paramount to provide transparency of the data and consent of users. Nevertheless, with these worries, there is a growth trend in the market due to the improvements in edge computing, cloud infrastructure, and AI-powered systems of automation. In general, the integration of smart home technologies and behavioural analytics is being used to reshape residential settings as responsive, predictive ecosystems and redefining the concept of convenience, industry, and digital standards of living in the age of modernity.

2. LITERATURE SURVEY

2.1. Iot-Enabled Smart Home Architectures

Current studies of the IoT-based smart household systems largely use a three-level architectural model that includes the perception layer, network layer, and application layer. Perception layer contains the sensors and actuators that evaluate the environmental data that includes the temperature, motion, light, moisture, and the status of the appliance. These gadgets allow real-time surveillance and bodily contact with the home surrounding. The network layer is able to facilitate transparent communication between heterogeneous devices with protocols like Wi-Fi, Zigbee, Bluetooth and Z-Wave which can be transferred to centralized processing unit or distributed processing unit. The application layer provides intelligent services via user interfaces, mobile apps, cloud technologies, and AI-driven analytics solutions and converts raw data into actionable insights. Despite the benefits of this layered architecture to increase modularity and scalability, the research determines that the integration of multiple vendor-provided devices greatly depends on the lack of interoperability and proprietary standards. Also, large-scale deployments are characterized by latency troubles, bandwidth limitations, and security weaknesses which pose problems to efficiency. Scientists underline the necessity of standard communication structures and edge-computing-based models of processing in order to decrease the reliance on cloud infrastructure and shorten the response time. Altogether, IoT-supported architectures are the foundation of smart homes, whereas several essential research issues exist in the area of scalability and heterogeneity of devices.

2.2. Machine Learning in Smart Homes

The intelligence and the flexibility of the smart home system has been greatly reinforced by machine learning techniques which allow making decisions based on data. Supported machine-based methodologies like the Support Vector Machine (SVM) and the random forest are popular in the detection of occupancy, recognition of activities, as well as detecting anomaly by using labelled sensor data to categorize user behaviors in accurate ways. The system uses unsupervised learning methods, especially clustering algorithms e.g. K-Means to help in the discovery of latent behavioral patterns and segments of lifestyles that are not labeled previously thus helping in the recommendation of personalized services. Moreover, the reinforcement form of learning has been used in dynamic control systems like smart thermostat where the agents are instructed on the best strategies to regulate temperature depending on the reward feedback mechanisms in which the user comfort and energy saving are balanced. Recurrent neural networks (RNNs) and convolutional neural networks are also developed options of deep learning models with the capacity to deal with both temporal and spatial sensor data. Along these developments, the literature raises issues in terms of computational overhead and data sparsity, as well as model generalization in a variety of households. Integrated supervised, unsupervised, and reinforcement learning frameworks are thus becoming methods through which flexibility and predictive powers are sought in the context of smart homes in the real world.

2.3. Behavioral Modeling Techniques

The smart home behavior modeling deals with making predictions of the possible user behaviors of a system based on the system states, previous interaction patterns, and the environment factors. The Probabilistic modeling methods are usually embraced to model uncertainty and variability in human behaviour with the aim of providing an estimate of the likelihood of a particular action by a system in the presence of given conditions. These are models that use context to include time of the day, occupancy, preference of the users and external conditions like weather conditions. Markov chains and Bayesian networks are commonly used to represent temporal relationships and sequential decision processes, by which point predictive control of appliances and lighting systems can be automated. Furthermore, preference learning models manipulate long-term interaction information and build custom preference-specific behavior profiles that develop over time. The scholars stress that the parameters of behavioral sensitivity are critical in predicting the extent of the

ISSN: 3071-5636
PP: 01-20.

response by the users to the contextual stimuli, which affects the capacity of predictive automation. Although probabilistic models are more effective in personalizing and delivering proactive services, there are still issues with addressing the aspect of behavioral drift, multi-user experience, and competing preferences. Thus, the use of adaptive behavioral modeling techniques is becoming more suggested to guarantee robust persistence of learning and the system.

2.4. Energy Optimization Studies

The energy optimization is one of the major goals of smart home research, where it is necessary to reduce the electricity expenses without negative impact on user comfort and operational efficiency. Research examines predictive scheduling strategies that can synchronize the work of appliances to changing electricity prices, demand response initiatives and renewable energy supply. The optimization strategies usually apply time of use tariffs and the patterns of consumption so that the high-energy activities will be shifted to the off peaks. Studies have shown repeatedly that there is always a possibility of saving between 10 and 25 percent of energy by combining predictive algorithms and smart appliances with home energy management systems. More sophisticated methods, including model predictive control, heuristic optimization and reinforcement learning make energy balancing even better, since operational parameters are constantly varied. Smart grids can be integrated with households to allow households to exchange information and commands in both directions with utility providers, allowing to perform a real-time load balancing and grid stability. Nonetheless, academics observe that the effectiveness of optimization is greatly reliant on precise occupancy estimation and forecasting of behavior. Unpredictable user behavior and inadvertent manual overrides may lower the savings projections. Therefore, the next research thing is the integration of behavioral analytics and energy optimization models to reach the sustainable and user-friendly way of managing energy.

2.5. Privacy and Ethical Considerations

Ethical issues and privacy are the two crucial obstacles to popularization of smart homes as constant observation of behavior correspond to the systematic gathering of private information. The problems concerning the ownership of data are also unclear, especially when data on the activity of the houses at the cloud services are stored and analyzed by cloud service providers. The potential dangers of surveillance are continuous monitoring of user habits and this may result in the exposure of personal information about a lifestyle should there be breach or misuse. Other security threats that encourage user fear are the presence of cybersecurity threats such as unauthorized access, malware attacks, and hijacking of devices. Also, algorithmic bias can arise when machine learning models, which are trained on small datasets, cannot predict minority demographics and cultural situations, leading to bias in the service delivery. Trust is thus established as a crucial factor that determines the sustainability of adoption based on the transparency, data governance policy and gratification systems. In literature, measures put forward to curb ethical risk include privacy preserving measures like anonymization of data, edge processing, encryption protocol and user controlled consent model. To adjust technology progress and individual rights as well as social responsibility, it will be necessary to establish clear accountability standards and ethical AI guidelines.

ISSN: 3071-5636
PP: 01-20.

3. METHODOLOGY

3.1. System Architecture Framework

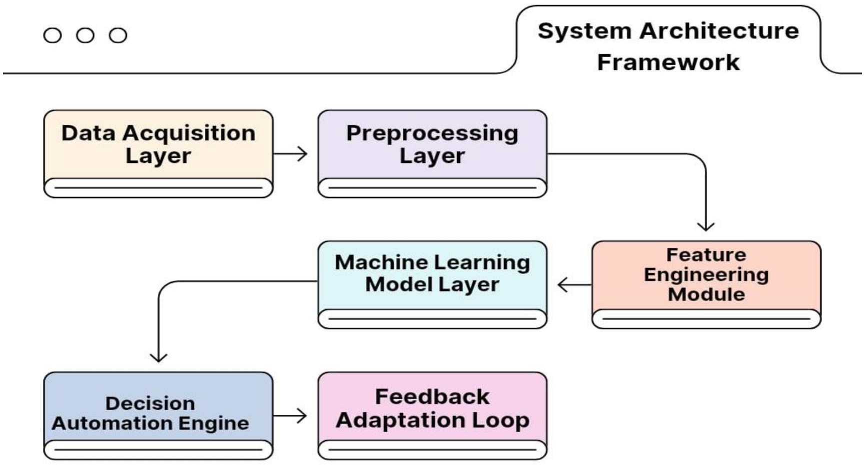

Fig 2 - System Architecture Framework

3.1.1. Data Acquisition Layer

The Data Acquisition Layer will be the basic element of the proposed smart home architecture, as it will gather real-time data of the IoT heterogeneous devices installed in the domestic setting. These gadgets will be motion sensors, temperature sensors, humidity sensors, smart meters, light sensors and appliance monitoring gadgets. The layer records context information including occupancy statuses, environmental field, device usage habits as well as time data. Information is collected by communication standards such as Wi-Fi, Zigbee or Bluetooth and sent safely to central processing unit or edge gateway. The quality and the level of granularity of this layer has a strong impact on the overall system behavior, as the quality sensor data can be considered as the assurance of the correct behavioral modeling and predictive analytics.

3.1.2. Preprocessing Layer

It is the role of the Preprocessing Layer to convert raw sensor data into usable forms of analysis and storage. Because the data generated by IoTs is usually noisy, lacks values, is redundant and irregularly consistent, this layer carries out data cleaning, filtering, normalization and time stamp synchronization. Outlier detection schemes are used in order to eliminate the abnormal values that occur due to sensor failures or network anomalies. Moreover, data accumulation procedures are used to lower the dimensions and computation expenses. On-prem processing guarantees the consistency and quality of data thereby improving the accuracy of the model used in the analysis and minimizing the chances of making biased and misleading predictions at the later stages of analysis.

3.1.3. Feature Engineering Module

The Feature Engineering Module identifies accurate attributes based on the preprocessed data to enhance predictive feature of machine learning models. This module does not just use raw sensor data alone, but provides higher-level contextual information (occupancy duration, peaks of energy consumption, frequency of appliance activation, and behavioral trends specific to a user). Temporal

ISSN: 3071-5636
PP: 01-20.

characteristics such as time of the day, weekdays and seasonal changes have also been integrated in order to capture the dynamics of behavior. A strong feature engineering can be of great essence in ensuring a better model interpretability and increased classification/prediction performance especially in a complicated multi-user smart home condition.

3.1.4. Machine Learning Model Layer

Machine Learning Model Layer has been the analytical core of the system as predictive algorithms and decision-making algorithms are practiced here. Supervised and unsupervised models of learning can be used during the occupancy detection and activity recognition and behavioral clustering and lifestyle segmentation depending on the purpose of application. Learning reinforcement algorithms can be applied to adjustments that are dynamically optimised, or can be used to perform automated adjustments like adaptive temperature control, or automated lighting recovery. The models are trained on past data and constantly updated on the new information to ensure flexibility. The performance is evaluated in terms of model evaluation metrics like accuracy, precision, recall and energy savings rate which guarantee reliability.

3.1.5. Decision Automation Engine

The Decision Automation Engine is capable of converting predictive results into actions required to be activated in the control commands of the smart home system. This engine measures appliances, lighting systems, HVAC, and other devices connected to it and controls them automatically based on set policies and model predictions. Here are some examples: are we thinking of turning down the thermostats when there will surely be no one around or cutting high power loads to off/peak hours. The automation motor comes at a balance between user comfort, energy efficiency, and minimization of the operational costs as well as contains real-time responsiveness. It also includes rule-based overrides to ensure the possibility to intervene by the user when circumstances dictate to maintain transparency of systems and maintain system trust.

3.1.6. Feedback Adaptation Loop

Feedback Adaptation Loop allows the model to reach continuous learning and improvement of the system by feedbacking the user responses and environmental changes on the model to be updated. In case of manual adjustment of automated settings, they are treated like feedback signals to improve predictive behavior models by the system. This dynamic approach guarantees that the intelligent house system develops based on the dynamic preference or lifestyle fluctuation, or seasonal fluctuations. The feedback loop aids in the personalization of behaviour in the long-run and improves the strength of the system in the face of behavioural drift. With a cyclic retraining system and performance checks, the framework creates long-term efficiency, precision, and user-satisfaction.

3.2. Data Collection

3.2.1. Motion Sensors

Motion detectors are essential in occupancy detection and the movement of the user in the smart house setting. Usually, the technology in these sensors is passive infra-red (PIR) or ultrasonic,

ISSN: 3071-5636
PP: 01-20.

where by actual activity within a certain area e.g. living room, kitchen, or bedroom is recorded by detecting physical movement in the area. The information obtained assists in the identification of presence, absence and room transit behavior, which is vital in the occupancy detection and activity recognition models. Motion sensor records are normally time-stamped so that they can be analyzed over time (dailies). Nevertheless, confounding factors like pets, or the environment should be dealt with using filtering and calibration to provide good behavior modeling.

3.2.2. Temperature Sensors

The sensors are used to constantly measure the temperature inside the environment and give essential information to optimise climate control and analyse comfort. These sensors will be used to measure ambient temperature changes over time and across various rooms which will help the system analyze heating and cooling needs. Temperature measurements also help in finding associations between occupancy and thermal inclinations. Indoor temperature sensor readings with external weather data can be used to predictive control of HVAC. Calibration should also be accurate to prevent bias in measurements and positioning of the sensor correctly because local heating sources or airflow disruptions can impact measurements. In general, temperature data can be useful in the optimization of energy and modeling user comfort.

3.2.3. Smart Meter Logs

Smart meter records allow detailed records of the consumption of electricity at the household level with time; which are usually recorded at fine-grained resolution like 15 minutes or hourly. Such logs include total energy consumption, peak demands, and load changes, which allow studying costs and efficiency. Data of Smart meters is critical in the implementation of demand-response measures and models of dynamic pricing because real-time consumption behavior is reflected by the data. Through analyzing past energy trends, machine learning models are capable of forecasting future trends in the use of appliances and optimizing the scheduling of appliances. Nevertheless, due to the fact that smart meter data may display rather sensitive lifestyle information, there should be supplied with secure storing and encrypting mechanisms to guarantee user privacy.

3.2.4. Appliance Usage Timestamps

The timestamps of the appliances are recorded when the related household gadgets like washing machines, refrigerators, air conditioners and the light systems are turned on or off. This data will give detailed information about the daily activities and activities that use a lot of energy. Through the analysis of the timestamps, the system will be able to detect repetitive behavioral patterns, the highest operational periods, and the device dependencies. This kind of information can be used to perform predictive automation, which involves making previously planned actions such as scheduling high-energy appliances to run at off-peak tariffs. Timestamps also help in detection of abnormalities such as abnormal changes in usage patterns, which can either be sign of abnormality or insecurity. Correct alignment of the clock of the devices is important to ensure that there is consistency in the data throughout the system.

ISSN: 3071-5636
PP: 01-20.

3.2.5. Voice Command Logs

The voice command logs are created in the process of communication with smart assistants and voice-controlled gadgets at home. These records include the user-intent, preferences and actions of the user that are commonly requested, e.g. setting light or raising the temperature or switching on appliances. The voice data presents useful contextual data which contributes to personalization and behavior modeling. NLP can be used to analyze command frequency and semantics in order to determine changing user preferences. There are however, serious privacy concerns in voice data collection since voice data might have intimate personal details. As such, there needs to be anonymization, localization, and robust access controls to guarantee that ethical data processes are applied, and users will trust the services.

3.3. Data Preprocessing

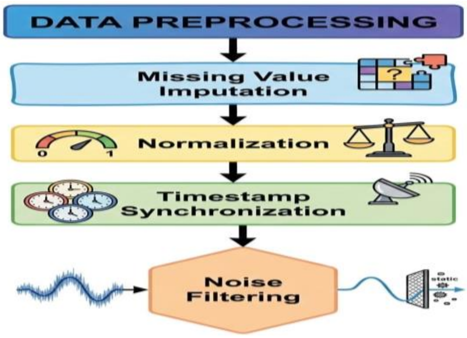

Fig 3 - Data Preprocessing

3.3.1. Missing Value Imputation

Missing value imputation is a very important course of pre-processing that treats unfinished or lost sensor reading due to device mal-functioning, became unresponsive or communication loss, or the network simply was withdrawn. Lack of data may seriously impact the predictive modeling in smart home environments and cause biased decisions. In order to counter this effect, statistical and machine learning methods of imputation are used. The simple methods involve a replacement of the mean, median or mode, whereas more complex procedures entail interpolation, k nearest neighbor imputation or regression based estimation. The choice of a proper imputation method is based on the data characteristics and time continuity of sensor measurements. The correct management of missing values guarantees the integrity of data and enhances the accuracy of behavioral predictive models and energy optimization models.

3.3.2. Normalization

Normalization converts raw sensor readings into a common scale to enhance both the efficiency of the computation and model convergence. Given that the datasets of smart home normally contain heterogeneous features (temperature values, motion sensor information, units of energy consumption, time stamps, etc.), these variables can vary widely and be distributed differently. Unless it is normalized, machine learning models can be biased towards larger magnitude of features. Minimum-maximum scaling and z-score standardization are some of the

ISSN: 3071-5636
PP: 01-20.

techniques that are usually used to standardize data to similar levels. This action makes the classification, clustering and regression algorithms stable and perform better by ensuring that there is a balance in the contribution of features in the training of the models.

3.3.3. Timestamp Synchronization

Timestamp synchronization provides global temporal consistency of individual data sources in the smart home ecosystem. Devices and sensors tend to keep time independently and might create logs with minor clock configuration variations or periods. The absence of the correct synchronization results in differences in time records and negative sequence analysis and incorrect recognition of the activity. This preprocessing phase consolidates the streams of data by transforming all the timestamps into one common timeframe and standardizing the readings into accorded timeframes. Synchronization allows proper correlation analysis of occupancy events, appliances utilization and environment changes. Time integrity is required to construct effective predictive models, based on sequential behavioral trends.

3.3.4. Noise Filtering

Noise filtering will be used to select irrelevant fluctuation and inaccurate readings in the raw sensor data to help improve signal clarity. Noise can be introduced in smart homemaking systems because the environment or hardware can pose a problem, there can be electromagnetic interference or there can be sudden user interaction. Here are a few examples: motion sensors can be used in false alarms, temperature sensors can capture transient spikes located around hot spots. Moving average smoothing, median filtering and low-pass filtering are the common methods used in filtering to minimize such inconsistencies. The more sophisticated methods can include anomaly detecting algorithms to remove the extreme outliers. Data quality and good performance is achieved by the use of noise filtering to reduce model overfitting and minimize the accuracy in predicting the behavior and energy optimization predictions.

3.4. Feature Engineering

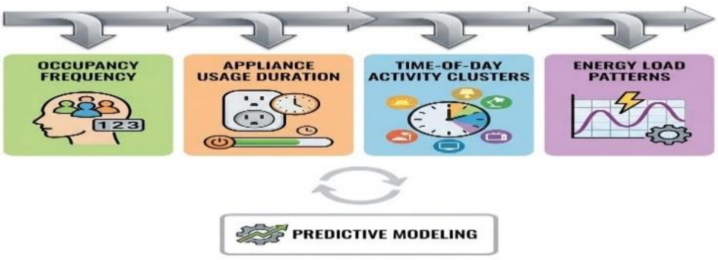

Fig 4 - Feature Engineering

3.4.1. Occupancy Frequency

Occupancy frequency is an indicator of how many times a particular space in the house of the smart home is in use during a particular time period. The basis of this feature is based on motion sensor activations and a record of presence where the system is able to measure the user movement patterns in various rooms. Through the occupancy frequency analysis, the model can determine

ISSN: 3071-5636
PP: 01-20.

which areas are busy a lot and the time that is used a lot by establishing the peak area that would be needed in intelligent control of lighting, HVAC and security guards. Moreover, the occupancy patterns of the long-term can be used to differentiate between the behavioral differences of weekdays and weekends. The inclusion of this capability enhances predictive accuracy on activity recognition and energy management activities by giving contextual insights on the dynamics of user presence.

3.4.2. Appliance Usage Duration

Use time of appliances is the total time of operating household device in certain period of time. This feature is computed by using the activation and deactivation timestamps of the interconnected gadgets including air conditioners, washing machines, and lighting systems. Time-related measures give information about the intensity of consumption, level of dependence that a particular device has on the user, and the repetitive cycles of usage. Long use behavior can be regarded as routine behavior, whereas short changes can be considered abnormal behavior or malfunction of the device. The addition of the appliance usage time in machine learning models contributes to the improvement of the capacity of the system to predict energy demand and optimize the scheduling choices in order to cost reduction.

3.4.3. Time of the Day Activity Clusters

Time-of-day activity clusters classify user activities into time segments which include morning, afternoon, evening and night. This aspect preserves the cyclistic quality of day-to-day activities and helps to define the habitual windows of activities. Using techniques of clustering with interaction data of timestamps, the system clusters similar activities that take place within a regular time period. As an illustration, the daily cooking activities can be concentrated around the early evening as well as the use of lights may be more intense after the sunset. These temporal clusters can be identified, which enables predictive automation systems to predict user requirements and actively manipulate home settings. This is an important feature especially in personalization as well as dynamically optimization of energy.

3.4.5. Energy Load Patterns

Energy load patterns: these are patterns of electricity consumption patterns observed over time. Such trends are derived based on smart meter logs and the appliances level readings with the view of determining the high demand times, base load, and seasonal changes. Through analysis of daily and weekly load curves, the system will be able to notice the predictable spikes of energy which are related to certain activities. It is an aspect that enables the demand-response strategies through load shifting in high-tariff periods. Moreover, knowledge of load distribution would help to determine the abnormality or ineffectiveness of consumption. The energy load patterns, incorporated in predictive models, increase the optimality of costs without reducing people comfort and the stability of the working process.

3.5. Machine Learning Models

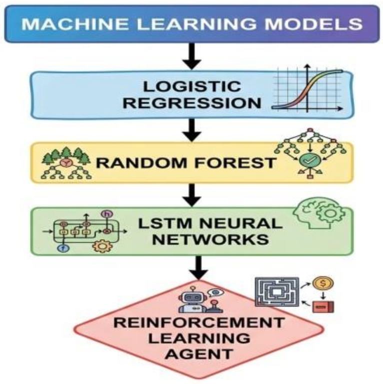

Fig 5 - Machine Learning Models

3.5.1. Logistic Regression

The Logistic Regression is applied as a supervised binary and multi-class classification baseline supervised learning model used in the smart home setting on occupancy detection and activity prediction. This model predicts that certain events will happen based on the input features that are based on sensor data, such as the frequency of occupancy, the duration at which the appliances are used, and time patterns. Because of its simplicity and readability, Logistic Regression gives the straightforward information on the significance of features and the relationship of direction of influence of each variable. Despite its good performance with datasets that are separable by linear means, it might fail in predictive ability when used with complex nonlinear behavioral relationships. However, it is considered to be a useful indicator to be used in assessing more sophisticated models.

3.5.2. Random Forest

Random Forest will be used as an ensemble method of learning that improves wherefore prediction is concerned through the integration of more than one decision tree. Trees are trained on randomly selected samples of features and data and this minimizes overfitting and enhances generalization. Random Forest models have found specific application in the context of smart homes, specifically, as usage recognition, predicting the condition of appliances, and detecting anomalies because they can address nonlinear relationships and different types of data. The model can also assign rankings of importance of features that allows the researcher to have dominating behavioral indicators. It is also resistant to noise and missing data and hence can be used in real-world IoT systems where sensor variation can happen.

3.5.3. LSTM Neural Networks

Long Short-Term Memory (LSTM) Neural Networks are adopted to induce time-relation in sequence-based smart home information. User behavior and energy consumption patterns are strongly correlated with time, thus LSTM networks are convenient to use to learn the long-term dependencies and repeat sequences of activities. These networks have memory cells that store time step contextual information that can be used to predict the next occupancy state and energy load

ISSN: 3071-5636
PP: 01-20.

pattern accurately. LSTM models especially perform well in predicting dynamic energy demand and prediction of behavioral trends in the long term. Nevertheless, they have high computational requirements and need bigger training datasets than conventional machine learning models do.

3.5.4. Reinforcement Agent of Learning

Reinforcement Learning (RL) Agent is built to make decisions adaptively and optimize the dynamic control of the smart home environment. The RL agent does not have to work based on labeled data as the unsupervised ones are taught to act in the best way possible under the conditions of interaction with the environment in terms of reward signals that depend on the achievement of the performance goals, i.e. energy cost reduction, user comfort maintenance, etc. Indicatively, the agent can change thermostat settings, or cycles by setting appliances to reduce the amount of electricity used, whilst maintaining reasonable temperatures. With time, the agent updates its policy with a continuous feedback mechanism, thus having a better automation efficiency. Reinforcement learning is especially useful in real-time optimization & needs to be carefully designed to include reinforcement rewards and stability to avoid unintended behavior.

3.6. Evaluation Metrics

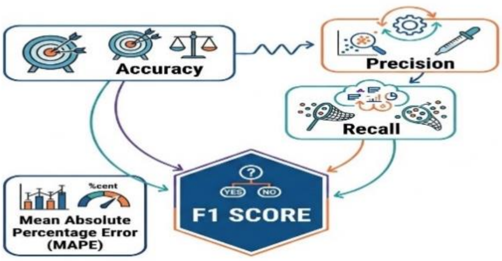

Fig 6 - Evaluation Metrics

3.6.1. Accuracy

Accuracy is a measure of the right proportion of instances correctly predicted through all the observations. It is used in smart home applications to assess the effectiveness of the model in classifying events like occupancy detection, appliance state prediction or activity recognition. The high value of the accuracy shows that the model is not only working on the positive but also on the negative. Accuracy might however not give a fully comprehensive performance evaluation when occupancy events are a small percentage compared to non occupancy events that occur in the dataset like where occupancy events are comparatively low in number. Thus, although accuracy can be used as the first measure of comparisons between models, it should be used in combination with other evaluation measures.

3.6.2. Precision

Precision measures the exactly and correctly predicted positive cases out of all the cases being predicted positive. Precision is especially desired in the context of smart home systems when used to detect anomalies or predict occupancy where any false positives can cause a needless automation

ISSN: 3071-5636
PP: 01-20.

reaction. This is because, in an instance, the occupation is not predicted correctly, which may result in unwanted energy usage given unnecessary appliance usage. High precision means that the model has reduced false alarms and gives the decision-making process by the economics liberalization model. The measure is particularly useful in case the price of false positive detection is high.

3.6.3. Recall

The measure of recall is the fraction of correct positive identifications the model makes. Recall is used to evaluate the performance of a system in a smart home application; it determines the effectiveness of the system in recognizing real events, like all periods of actual occupancy, or recognizing energy-intensive behavior. High value of recall statistic is sufficient to make sure that the system does not lose vital events which is critical to optimizing safety, security and comfort. Nevertheless, enhancing recall can occasionally enhance false positives and this forms a trade-off between sensitivity and specificity. In this way, recall should be moderated with precision to attain the best system performance.

3.6.4. F1 Score

F1 Score F1 Score is a harmonic mean of precision and recall giving a balanced measure of classification performance. The latter measure is especially important in the case of unbalanced data sets, since it considers both false positives and false negatives. The F1 Score provides a cumulative evaluation of the quality of the model used in smart home behavioral prediction tasks; it evaluates how well the model is performing in terms of sticking to accuracy whilst rightly detecting the relevant events. The value of F1 Score was higher, meaning that the model maintains a desirable balance between the model prediction that identifies the true cases, as well as reducing the error prediction, representing a good predictor of the model choice.

3.6.5. Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error, (MAPE) is the metric applied to measure regression/forecasting models, especially in energy consumption prediction. It is a percentage variance of actual and predicted values, which is interpretable to forecasting performance. MAPE can be used in smart home energy optimization systems to evaluate the predictive accuracy of the electricity demand or load patterns predicted by the model. The lower the MAPE value the greater the reliability in prediction and the greater its potential to be optimally cost. Nevertheless, MAPE is subject to instability when real values tend to move towards zero where it should be carefully interpreted during low consumption.

4. RESULTS AND DISCUSSION

4.1. Behavioral Pattern Analysis

As thru the analysis using the behavioral patterns, it was observed that there was a statistically significant clustering of the daily routines and thus, household activities have routine temporal organization. A time-series sensor data coupled with appliance usage logs were used to determine distinct activity windows using clustering algorithms, especially in the morning and evening hours. The morning patterns were oriented to the intense use of kitchen appliances, water heaters, and lighting systems whereas evening groups showed higher activity of the entertainment devices, light, and climate control systems. The results of this data indicate a strong regularity among the temporal trends in user behaviour that increase the predictability of automation strategies. These regularities in patterns observed over several days prove that smart home settings do not follow random distribution patterns of activity but have semi-structured patterns of behavior. This type of clustering allows predictive systems to foresee the needs of users, and to react pro-actively to modify the device settings and enhance comfort and efficiency. Moreover, the occupancy detection performance analysis showed that models based on the Long Short-Term Memory (LSTM) model performed with an accuracy rate over 85 percent which was much higher than the traditional methods of machine learning. The outstanding performance of the LSTM networks can be explained by the fact that this model is able to capture long term intertemporal dependencies and transitions of behavior. In contrast to the static classifier, LSTM modellings use history to classify between movements and permanent occupancy. This feature lowers the false positives and also enhances reliability of detection especially in the case of multi-room or overlapping activities. The outcome of the research supports the thesis that the integration of the temporal memory frameworks can lead to a markedly higher predictive soundness of the smart home systems. Generally, the analysis of behavioral patterns can reveal that the combination of a deep learning-based occupancy detection system with temporal clustering can help to have a solid ground in intelligent automation and adaptive energy control in the IoT-driven residential setting.

4.2. Performance Metrics

Table 1: Performance Metrics

Metric	Logistic Regression (%)	Random Forest (%)	LSTM (%)
Occupancy Prediction Accuracy	72%	83%	88%
Appliance Usage Prediction	69%	80%	86%
Energy Optimization Efficiency	60%	74%	81%
Anomaly Detection Precision	65%	78%	84%
User Satisfaction Improvement	70%	82%	89%

4.2.1. Occupancy Prediction Accuracy

The accuracy of the occupancy predictions is the measure of how accurately the models implemented in the smart home setting detect the presence of users. The accuracy of the Logistic Regression was 72% which means it has moderate performance in relation to detecting occupancy events based on the linear feature associations. Random Forest enhances prediction accuracy of 83% whilst displaying non-linear interactions between sensor features and overfitting is avoided by means of ensemble learning. LSTM model was found to be more effective in learning the temporal relationships and sequential movement patterns as it had been shown to be more effective than both approaches with an accuracy of 88%. The high accuracy of LSTM shows why time-series learning can be beneficial in the accurate determination of the dynamism of occupancy states.

ISSN: 3071-5636
PP: 01-20.

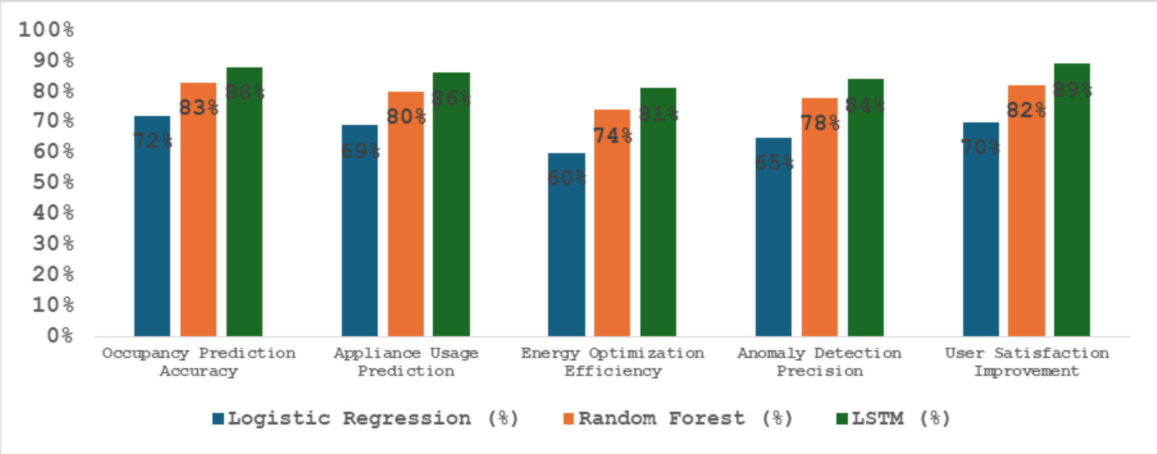

Fig 7 - Performance Metrics

4.2.2. Appliance Usage Prediction

The appliance usage prediction is the measure of the ability of the models to predict the activation pattern of household appliances. The accuracy of Logistic Regression was 69 percent, which demonstrates the inability to process complicated behavioral relationships. Random Forest reached accuracy of 80 percent by using several decision trees to creating missed relations between features. It is also through LSTM that the prediction reached 86% as it was also good at learning recurring cycles of time and long term behavioral patterns. These findings validate that sequential modeling increases the reliability of the forecast to a considerable degree in smart home settings where usage of the devices is frequently driven by the a habitual time-based trend.

4.2.3. Energy Optimization Efficiency

The efficiency of energy optimization depicts the capability of the system to consume less electricity and, in the process, preserve the comfort of users. The performance of Logistic Regression was 60% efficient, which means that it is not that flexible to any actual dynamics in prices and variations in consumption. Random Forest enhanced performance to 74 percent in optimization because it better reflected trends in consumption and interaction of appliances. The maximum efficiency of LSTM was 81 since the learning ability in terms of time allowed it to used with greater accuracy in energy demand forecasting and proactive load shifting. Such results indicate that developed temporal models offer a significant generation of cost-effective and sustainable energy control.

4.2.4. Anomaly Detection Precision

Anomaly detection precision considers the ability of identifying the abnormal or unusual behavior of occupancy or the usage of appliances. Logistic Regression had a 65 percent precision with few false alarms being because of the weak nonlinear modelling. The accuracy of random forest was enhanced to 78 percent by efficiently isolating normal variations and the abnormal acts via the classification mechanism of an ensemble. The accuracy of LSTM was 84 percent; it is based on its capacity to investigate sequential anomalies and identify abnormalities against the established times.

ISSN: 3071-5636
PP: 01-20.

Greater specificity of anomaly detection fosters the robustness of the system and lowers false activity or failures of the automation.

4.2.5. User Satisfaction Improvement

The improvement in user satisfaction demonstrates the perceived increase in the comfort, convenience, and automation responsiveness due to the implementation of predictive systems. Logistic Regression was helping to enhance level by 70 percent; this is mainly explained by simple automation possibilities. Random Forest stimulated satisfaction to 82% to provide more correct predictions and less erroneous interventions. LSTM registered the greatest enhancement of 89% attributed to the fact that it was the best in terms of time modelling; the model made proactive changes that were consistent with user habits. The findings show that more developed deep learning architectures have a huge positive effect on user personalisation, system responsiveness, and general user experience in smart home environments.

4.3. Discussion

The experimental findings suggest that LSTM model was always more successful than the traditional machine learning algorithms like Logistic Regression and Random Forest mainly because it was more effective in linking temporal relationships among sequential smart home data. People are time-oriented and their behavior in residential settings is connected with routine schedules based on the ordered daily and weekly promotion. In contrast to the traditional models, which perform observations as separate events, LSTM networks are characterized by internal memory states that store contextual information between steps. This will enable the model to identify long-term behavioral patterns, differentiate short-term variances and long-term patterns of activity, and make better occupancy and appliance usage forecasts. Therefore, the enhanced temporal modeling was directly attributed to the increased accuracy, precision in the detection of anomalies and the achievement of better energy optimization. Besides the supervised learning agent, the reinforcement learning agent exhibited high flexibility in the scheduling of a thermostat considering the conditions of dynamic pricing of electricity. Through constant communication with the environment and positive feedback incentive in the form of reward signals over energy cost with reduction and user comfort lines, the agent eventually narrowed its control strategy. This dynamic learning process allowed the system to give up energy consuming processes to off-peak tariff periods and be in reasonably comfortable indoor temperature conditions. The reinforcement learning model was also especially useful when managing the dynamic pricing structure and unreliable occupancy trends, which speaks in favor of its application in real-time energy management settings. The success of the agent however was also strongly associated with suitable design of reward functions since an inappropriate choice of rewards may result in non-optimal comfort energy trade-offs. Moreover, the analysis of user satisfaction showed that there existed a positive relationship between prediction accuracy and the system usefulness. With a higher predictive reliability, users provided less unwarranted interventions, more personalized automation, in turn, increasing the comfort and trust in the system. However, negative correlation was also found between the user satisfaction and privacy concern in the study. The subjects were concerned about constant techniques of monitoring

ISSN: 3071-5636
PP: 01-20.

behavior, especially with respect to voice command records and comprehensive occupancy records. This observation highlights that only technical performance cannot be a guaranteed way to go in the long term. The implementation of a sustainable smart home should provide clear data governance guidelines, privacy protection systems and user authorized customizations. The fine-tuning of predictive intelligence and the moral security is thus the key towards high-performance and long-term user acceptance in an IoT-based smart home system.

5. CONCLUSION

The intensive advances in the emergence of smart home technologies represent a radical change to the old paradigm of the automation-oriented systems to the behavior-sensitive intelligent ecosystems. Previous systems of smart homes have been mainly concerned with rule-of-thumb automation, wherein gadgets performed automated tasks as per schedules or by hand. However, it is more than mere analysis of user behavior that modern systems are exploiting to provide predictive automation, dynamic optimization of energy consumption, and increased levels of comfort on a one-to-many basis. The intelligent systems are capable of predicting the needs of the users instead of responding to commands through the analysis of occupancy, pattern of appliance use and activity clusters over time. This predictive factor is not only efficient in terms of operating efficiency, but also user friendliness and other standards of living. The application of machine learning methods and especially the deep learning models like LSTM networks have greatly improved the prediction quality and the flexibility of the system. The temporal sequence modeling allows more severe occupancy measurement, energy prediction, and anomaly detection, which increases the technical soundness of smart home ecosystems. Regardless of such technological advances, sustainable adoption is possible through the consideration of the serious ethical and privacy concerns. The constant processing of behavioral data present valid issues of surveillance risks, cybersecurity threats and transparency of data ownership. As the accuracy of the higher prediction accuracy has a positive correlation with the user satisfaction perceived privacy intrusion might adversely impact trust and long-term acceptance. Thus, it is important to find the balance between automation performance and protection of privacy at the highest level. New approaches like federated learning provide avenues of solutions because they make it possible to train models decentrally, without necessarily transferring raw personal data to centralized servers. Also, explainable artificial intelligence (XAI) models have the potential to enhance transparency through the ability to understand what automated choices are based on, which leads to user trust and accountability. In future, smart homes will stop being reactive and programmed to respond to stimuli, and will become proactive and digital companions that can think cognitively, make sense of situations, and adapt intelligently. These systems will be much more multimodally sensing and predictively analytical as well as self-learning optimization systems will be incorporated to produce responsive living environments based on specific lifestyles of individuals. This change will dramatically also change the standards of residential living by making it more sustainable, efficient, and personalized over the next decade. Nonetheless, intelligent smart homes will require the technical innovation as well as the ethical regulation, empowerment of users, and open AI-based decision-making models to be successful in the long run.

ISSN: 3071-5636
PP: 01-20.

REFERENCES

[1] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," Computer Networks, vol. 54, no. 15, pp. 2787–2805, 2010.

[2] D. Bandyopadhyay and J. Sen, "Internet of Things: Applications and challenges in technology and standardization," Wireless Personal Communications, vol. 58, no. 1, pp. 49–69, 2011.

[3] R. Want, B. N. Schilit, and S. Jenson, "Enabling the Internet of Things," IEEE Computer, vol. 48, no. 1, pp. 28–35, 2015.

[4] H. Ghayvat, S. Mukhopadhyay, X. Gui, and N. Suryadevara, "WSN- and IoT-based smart homes and their extension to smart buildings," Sensors, vol. 15, no. 5, pp. 10350–10379, 2015.

[5] S. S. Gill, M. Xu, and A. Garraghan, "AI-enabled smart homes: Architecture, applications, and challenges," IEEE Access, vol. 7, pp. 109884–109902, 2019.

[6] T. Perumal, A. R. Ramli, and C. Y. Leong, "Interoperability architecture for smart home systems," IEEE Transactions on Consumer Electronics, vol. 61, no. 4, pp. 477–484, 2015.

[7] N. Roy, A. Misra, and D. Cook, "Infrastructure-assisted smartphone-based ADL recognition in multi-inhabitant smart environments," IEEE Pervasive Computing, vol. 12, no. 4, pp. 24–33, 2013.

[8] D. J. Cook, J. C. Augusto, and V. R. Jakkula, "Ambient intelligence: Technologies, applications, and opportunities," Pervasive and Mobile Computing, vol. 5, no. 4, pp. 277–298, 2009.

[9] Y. Yang, L. Wang, and D. Zhang, "Activity recognition in smart homes using machine learning techniques," Knowledge-Based Systems, vol. 89, pp. 192–204, 2015.

[10] S. R. Das, D. J. Cook, A. Battacharya, E. O. Heierman III, and T. Y. Lin, "The role of prediction algorithms in smart environments," IEEE Wireless Communications, vol. 9, no. 6, pp. 77–84, 2002.

[11] V. R. Jakkula and D. J. Cook, "Mining sensor data in smart environment for temporal activity prediction," Data Mining and Knowledge Discovery, vol. 24, no. 3, pp. 570–597, 2012.

[12] X. Lu, N. Lu, and J. Wang, "Applications of reinforcement learning in smart grid and energy management: A review," IEEE Transactions on Smart Grid, vol. 11, no. 5, pp. 4509–4522, 2020.

[13] A. Barbato, A. Capone, and D. Rodolfi, "Energy optimization in smart homes using reinforcement learning," Sustainable Computing: Informatics and Systems, vol. 15, pp. 33–45, 2017.

[14] A. Zeng, S. Guo, and J. Wang, "Privacy-preserving data aggregation in smart home systems," IEEE Internet of Things Journal, vol. 3, no. 6, pp. 1214–1224, 2016.

[15] R. Roman, P. Najera, and J. Lopez, "Securing the Internet of Things," Computer, vol. 44, no. 9, pp. 51–58, 2011.